Final Technical Report
NASA Grant NAG 5-1211
Statistical Algorithms for Handling
Low Count Per Bin Datasets

John A. Nousek

April 10, 1991

This report describes the conclusion of activities at Penn State University supported by NASA Grant NAG5-1211, "Statistical Algorithms for Handling Low Count Per Bin Datasets". This report covers the period from May, 1990 through April, 1991.

The purpose of this grant was to develop and test algorithms for fitting HEAO-2 Imaging Proportional Counter data in cases where the total number of counts per spectrum was small. The proposal was made to NASA through the Astrophysics Data Program NRA, and was accepted and funded for one year starting in July, 1989.

The proposal was based on work described in Nousek and Shue (1989), which tested the effect of applying χ^2 fitting to simulated datasets computed from an ideal distribution, as a function of the number of counts in the datasets. It was shown that the inferred 'best fit' parameters systematically differed from the original ideal ones. Nousek and Shue explored an alternative statistic proposed by Cash (1979), called the 'C' statistic, and found that it was free of bias.

Based on these results the proposal which was funded through this grant was to perform similar testing on ideal simulated spectra from the HEAO-2 IPC instrument. If a bias were detected, the 'C' statistic would be applied to the IPC data, and similarly tested. If the 'C' results were found to be bias free the fitting algorithms would be provided to the HEAO-2 analysis center at the Smithsonian Astrophysical Observatory in Cambridge, Massachusetts.

The activity supported by this grant consisted of the following. Dr. Rick Harnden of the Smithsonian provided the IPC response matrices on the Principal Investigator's request in July, 1989. Mr. Christopher Frye, a senior undergraduate Honors student (and University Scholar) accepted the IPC testing project for his undergraduate thesis project in August, 1989.

In the process of completing his undergraduate thesis in August, 1990, Mr. Frye successfully modified the code used by Nousek and Shue to incorporate the IPC response matrix and effective area, and tested the bias introduced by the χ^2 analysis. His conclusions were that the bias detected in the ideal case by Nousek and Shue is indeed present in the Einstein IPC case, but (at least for power law spectra of index -2) the size of the bias is roughly half that of the ideal case. Frye speculates that the smaller bias is due to a distribution of the bins with zero counts across the spectrum, rather than the concentration at the high energy end seen in the ideal detector case.

After Mr. Frye's graduation the C-statistic and χ^2 fitting code was restructured into a form which makes it usable in more general cases. This work was performed by Ms. Vida Farwana under the direction of Dr. John Nousek. Instructions covering the use of this software and a sample problem, plus a general description of the C-statistic methodology has been prepared, and a copy appended to this report.

The project will be completed on April 30, 1991. The final month's activities will consist of beta testing of the software by Mr. Greg Berthiaume, and distribution of the manual and software to Drs. Steve Murray and F. Rick Harnden at the Smithsonian Astrophysical Observatory and Drs. Nick White and Keith Arnaud of the High Energy Astrophysics Science Archival Research Center at Goddard Space Flight Center. During this period we would be prepared to also distribute copies to any sites designated by the NASA Science Operations Branch.

REFERENCES

1. Cash, W. 1979, *Ap. J.,* **228**, 939.

2. Nousek, J., and Shue, D. 1989, *Ap. J.*, **342**, 1207.